Solving Imperfect Recall Games

(Doctoral Consortium)

Jiří Čermák
Department of Computer Science
Czech Technical University in Prague
jiri.cermak@agents.fel.cvut.cz

ABSTRACT

Imperfect recall games remain an unexplored part of the game theory, even though recent results in abstraction algorithms show that the imperfect recall might be the key to solving the immense games found in the real world efficiently. Our research objective is to develop the first algorithm capable of outperforming the state-of-the-art solvers applied directly to perfect recall games, by creating and solving its imperfect recall abstraction. To achieve this, we need to answer a set of open questions considering the imperfect recall games. Namely (1) what properties allow the existence of Nash equilibrium in behavioral strategies in imperfect recall games and (2) what information are the players allowed to forget to guarantee that the resulting imperfect recall game can be used to solve the original game with a bounded error.

1. INTRODUCTION

Game theory is a mathematical model of strategic interaction between agents. It is a descriptive theory defining conditions for strategies of players to form an equilibrium. The most famous equilibrium, where no rational player wants to deviate from the prescribed strategy, is called a *Nash equilibrium* (NE). Simple one-shot problems are represented as *normal-form games*. In this work, however, we focus on more general situations, where we face finite sequential decisions including partial observability and a stochastic environment. For these situations, *extensive-form games* (EFGs) are the most suitable representation. EFG is represented as a game tree with nodes corresponding to the game states and edges representing actions available to players. The partial observability is represented by grouping states indistinguishable for the player making the decision to information sets. More specifically, we focus on *imperfect recall games* where the structure of partial information forces players to forget their moves or the information available to them in the past.

Recent advancements in scalability of algorithms for solving EFGs has been primarily driven by the research around the Annual Computer Poker Competition[1] and has led to solving heads-up limit texas hold'em poker [2]. Most of the algorithms for solving EFGs assume that players remember all the information gained during the game [15, 17] – a property denoted as a *perfect recall*. The size of a *behavioral strategy* (a randomized selection of an action to play in each information set) grows exponentially with the number of moves in the game due to the perfect memory. One approach for solving large perfect recall EFGs is thus to create a small abstracted game where certain information sets are merged, solve this abstracted game, and translate the strategy into the original game (e.g., see [5, 10]). However, to achieve sufficient space reductions the assumption of perfect recall might need to be violated in the abstracted game resulting in *imperfect recall*.

There are fundamental difficulties when solving imperfect recall games. The NE in behavioral strategies does not have to exist even in zero-sum games [16]. This is caused by a different descriptive power of behavioral and *mixed strategy* representation in imperfect recall games [12] (mixed strategies represents the behavior using a probability distribution over the whole deterministic assignments of one action to every information set of the player). The original proof of the existence of Nash equilibrium for finite games considers mixed strategies only [14]. However, the concept of mixed strategies is not directly applicable to games with imperfect recall, as it allows players to condition their actions on information hidden by the rules of the game. Furthermore, only behavioral strategies can fully benefit from the reduced number of decision points in the imperfect recall abstraction. For this reason, behavioral strategies are used in imperfect recall games [12]. Additionally, from the computational perspective, even to check whether an imperfect recall game has a NE in behavioral strategies is shown to be NP-hard [6].

An alternative to seeking NE is to find the behavioral strategy that maximizes the worst case expected outcome of a game—*a maxmin strategy*, which is guaranteed to exist in imperfect recall games. Maxmin strategies may require irrational numbers even when the input uses rational numbers in imperfect recall games [9]. Hence, computing optimal maxmin strategies has fundamental difficulties, and approximating maxmin strategies is equivalent to solving imperfect recall games. Also deciding whether a player with imperfect recall can guarantee at least a given value in an imperfect recall game is an NP-hard problem [9].

2. RELATED WORK

The fundamental difficulty of strategy representation and the difference between behavioral and mixed strategies in imperfect recall games was first discussed by Kuhn [12]. An

[1] http://www.computerpokercompetition.org/

Appears in: *Proc. of the 16th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2017), S. Das, E. Durfee, K. Larson, M. Winikoff (eds.), May 8–12, 2017, São Paulo, Brazil.*

example showing that imperfect recall games need not have Nash equilibrium in behavioral strategies, which builds on the difference of mixed and behavioral strategies described by Kuhn, was provided by Wichardt [16].

Kaneko et. al. [7] study the additional information introduced when using mixed strategies in imperfect recall games. They introduce the notion of A-loss recall games, which form a subset of imperfect recall games, where every loss of a memory of player can be tracked back to losing information about his actions. They show that in A-loss recall games the use of mixed strategies completely compensates the information hidden to players due to imperfect recall. In the follow-up work by Kline [8] it is shown that in A-loss recall games any strategy is ex-ante optimal if and only if players do not want to deviate during the actual playthrough. This property is called *time consistency*. Similarly, in parallel work Bonano [1] distinguishes the difference between imperfect recall caused by forgetting information one had previously available, and action one has previously played.

There is only a limited amount of work focused on finding the optimal strategies in imperfect recall games. Restricted classes of imperfect recall games were described, where the Counterfactual regret minimization [17] is guaranteed to converge to Nash equilibrium strategies [13, 11].

3. RESEARCH OBJECTIVES

Our goal is to devise an algorithm which will automatically detect the least relevant information players remember in a game and create a small imperfect recall abstraction by removing it. The algorithm will then exploit the reduced size of the abstraction and efficiently solve it while guaranteeing a bounded error in the original game.

We first explore the classes of imperfect recall abstractions which can be used to find rational behavior in the original unabstracted game. The result presented in [4] shows that certain subclasses of A-loss recall games have this property. Furthermore, we are interested in the computational complexity of solving the imperfect recall abstraction; hence this work summarizes the known complexity results concerning imperfect recall games and adds several previously unknown complexity results of solving A-loss recall games.

We build on the results in [4] and provide the first algorithm for finding maxmin strategies of a given imperfect recall game where the minimizing player has A-loss recall [3]. This algorithm is based on solving a relaxation of a bilinear mathematical program and incremental strategy generation.

Finally, as a future work, we plan to devise an algorithm, which given a perfect recall game creates its coarse imperfect recall abstraction. The algorithm will then simultaneously solve the abstraction and refine the information set structure based on the possible behavior of the players in such a way that the reached solution has a bounded error in the original perfect recall game.

4. ACKNOWLEDGMENTS

This research was supported by the Czech Science Foundation (grant no. 15-23235S), and by the Grant Agency of the Czech Technical University in Prague, grant No. SGS16/235/OHK3/3T/13. Computational resources were provided by the CESNET LM2015042 and the CERIT Scientific Cloud LM2015085, provided under the programme "Projects of Large Research, Development, and Innovations Infrastructures."

REFERENCES

[1] G. Bonanno. Memory and perfect recall in extensive games. *Games and Economic Behavior*, 47(2):237–256, 2004.

[2] M. Bowling, N. Burch, M. Johanson, and O. Tammelin. Heads-up limit hold'em poker is solved. *Science*, 347(6218):145–149, 2015.

[3] J. Cermak, B. Bosansky, and M. Pechoucek. Combining incremental strategy generation and branch and bound search for computing maxmin strategies in imperfect recall games. In *AAMAS*, 2017.

[4] J. Cermak, B. Bosansky, and M. Pechoucek. Towards solving imperfect recall games. In *CPW*, 2017.

[5] A. Gilpin and T. Sandholm. Lossless Abstraction of Imperfect Information Games. *Journal of the ACM (JACM)*, 54(5):25, 2007.

[6] K. A. Hansen, P. B. Miltersen, and T. B. Sørensen. Finding Equilibria in Games of no Chance. In *Computing and Combinatorics*, pages 274–284. Springer, 2007.

[7] M. Kaneko and J. J. Kline. Behavior Strategies, Mixed Strategies and Perfect Recall. *International Journal of Game Theory*, 24:127–145, 1995.

[8] J. J. Kline. Minimum Memory for Equivalence between Ex Ante Optimality and Time-Consistency. *Games and Economic Behavior*, 38:278–305, 2002.

[9] D. Koller and N. Megiddo. The Complexity of Two-person Zero-sum Games in Extensive Form. *Games and Economic Behavior*, 4:528–552, 1992.

[10] C. Kroer and T. Sandholm. Extensive-Form Game Abstraction with Bounds. In *EC*. ACM, 2014.

[11] C. Kroer and T. Sandholm. Imperfect-recall abstractions with bounds in games. In *EC*. ACM, 2016.

[12] H. W. Kuhn. Extensive Games and the Problem of Information. *Annals of Mathematics Studies*, 1953.

[13] M. Lanctot, N. Burch, M. Zinkevich, M. Bowling, and R. G. Gibson. No-regret learning in extensive-form games with imperfect recall. In *ICML*, 2012.

[14] J. F. Nash. Equilibrium Points in n-person Games. *Proc. Nat. Acad. Sci. USA*, 36(1):48–49, 1950.

[15] B. von Stengel. Efficient Computation of Behavior Strategies. *Games and Economic Behavior*, 14:220–246, 1996.

[16] P. C. Wichardt. Existence of Nash Equilibria in Finite Extensive Form Games with Imperfect Recall: A Counterexample. *Games and Economic Behavior*, 63(1):366–369, 2008.

[17] M. Zinkevich, M. Johanson, M. Bowling, and C. Piccione. Regret Minimization in Games with Incomplete Information. *NIPS*, 2008.